END

DATE
FILMED
6-77

REPORT DOCUMENTATION PAGE

1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
AFOSR - TR- 77- 0565		

4. TITLE (and Subtitle)	5. TYPE OF REPORT & PERIOD COVERED
MULTISTAGE ESTIMATION OF BIAS STATES IN LINEAR SYSTEMS.	Interim rept.
	6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)	8. CONTRACT OR GRANT NUMBER(s)
J. M. Mendel H. D. Washburn	AFOSR 73-2797-75

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
University of Southern California Department of Electrical Engineering Los Angeles, California 90007	61102F 2304/A1

11. CONTROLLING OFFICE NAME AND ADDRESS	12. REPORT DATE
Air Force Office of Scientific Research/NM Bolling AFB DC 20332	1977
	13. NUMBER OF PAGES
	27

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report)
	UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This paper provides an alternate, constructive derivation of Friedland's method (Ref 1) for recursive bias filtering; and, extends his method to the case where we may wish to increase (or decrease) the number of biases. We show that it is possible to add (or delete) bias states in such a manner that previously computed quantities can be used to obtain new estimates of the dynamical state vector and the now-larger bias vector. Adding (or deleting) bias states is important when, for example, the bias states are used to model constant, but unknown, over
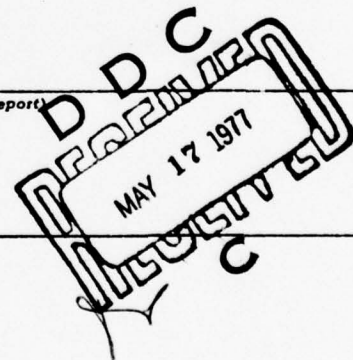
Block 20 - continued

instrumentation error sources, of which there can be a large number.

(2)

MULTISTAGE ESTIMATION OF BIAS STATES
IN LINEAR SYSTEMS*

by

J. M. Mendel and H. D. Washburn

Department of Electrical Engineering
University of Southern California
Los Angeles, California 90007

ABSTRACT

This paper provides an alternate, constructive derivation of Friedland's method (Ref. 1) for recursive bias filtering; and, extends his method to the case where we may wish to increase (or decrease) the number of biases. We show that it is possible to add (or delete) bias states in such a manner that previously computed quantities can be used to obtain new estimates of the dynamical state vector and the now-larger bias vector. Adding (or deleting) bias states is important when, for example, the bias states are used to model constant, but unknown, instrumentation error sources, of which there can be a large number.

I. INTRODUCTION

We shall be interested in estimating states, $\underline{x}$ and $\underline{b}$, for the following (continuous- or discrete-time) system, which is assumed to be the actual system or truth model:

$$\mathscr{A}_0 \begin{cases} & O(\underline{x}) = A\underline{x} + B\underline{b} + \underline{\xi} \quad ; \quad \underline{x}(0) \tag{1} \\ \\ \text{or} & \dot{\underline{b}} = \underline{0} \quad ; \quad \underline{b}(0) = \underline{b} \tag{2a} \\ \\ & \underline{b}(k+1) = \underline{b}(k) = \underline{b} \tag{2b} \\ \text{and} \\ & \underline{y}(\cdot \mid \mathscr{A}_0) = H\underline{x} + C\underline{b} + \underline{\eta} \tag{3} \end{cases}$$

In these equations, we use operator $O(\underline{x})$ to denote either $\dot{\underline{x}}(t)$ or $\underline{x}(k+1)$, and we do not show the explicit dependence of vector and/or matrix quantities on time. In this manner, we are able to present results for continuous-time and discrete-time systems simultaneously. The conditioning notation on the measurement is to remind us of the model which is associated with the measurement; $\underline{y}(\cdot \mid \mathscr{A}_0)$ will be referred to as the actual measurement. Models other than $\mathscr{A}_0$ will be introduced below. In $\mathscr{A}_0$, $\underline{x} \in R^n$, $A \in R^{n \times n}$, $\underline{b} \in R^r$, $B \in R^{n \times r}$, $\underline{\xi} \in R^n$, $\underline{y} \in R^s$, $H \in R^{s \times n}$, $C \in R^{s \times r}$, and $\underline{\eta} \in R^s$; and, $\underline{\xi}(\cdot)$ and $\underline{\eta}(\cdot)$ are gaussian white noise processes for which

$$E\{\underline{\xi}\} = \underline{0} \qquad \text{and} \qquad E\{\underline{\eta}\} = \underline{0} \tag{4}$$

$$E\{\underline{\xi}(t)\,\underline{\xi}'(\tau)\} = Q\,\delta(t-\tau) \quad \text{or} \quad E\{\underline{\xi}(k)\,\underline{\xi}'(\ell)\} = Q\,\delta_{k\ell} \tag{5}$$

$$E\{\underline{\eta}(t)\,\underline{\eta}'(\tau)\} = R\,\delta(t-\tau) \quad \text{or} \quad E\{\underline{\eta}(k)\,\underline{\eta}'(\ell)\} = R\,\delta_{k\ell} \tag{6}$$

and *

$$\underline{\xi} \perp \underline{\eta} \qquad (7)$$

Friedland (Ref. 1) describes a very interesting and practical procedure for estimating the dynamical states, $\underline{x}$, and the constant, but unknown bias states, $\underline{b}$, for system $\mathscr{A}_0$. He shows that the optimum estimate of $\underline{x}$, $\hat{\underline{x}}$, and its associated error covariance matrix, P_x, can be expressed as

$$\hat{\underline{x}} = \tilde{\underline{x}} + V_x \hat{\underline{b}} \qquad (8)$$

and

$$P_x = P_{x_2} + V_x P_b V_x' \qquad (9)$$

where $\tilde{\underline{x}}$ is the bias-free estimate of $\underline{x}$, computed as if no biases were present, $\hat{\underline{b}}$ is the optimum estimate of the bias, and V_x is a matrix which blends the estimates $\tilde{\underline{x}}$ and $\hat{\underline{b}}$ together to give $\tilde{\underline{x}}$, the bias corrected estimate of $\underline{x}$. Matrices P_b and P_{x_2} are defined in Section II.

In Friedland's derivation, he augments the bias states to the dynamical states and shows that the resulting Riccati equation decomposes so that $\hat{\underline{x}}$ is computed via Eq. (8). This derivation is algebraic in nature, and unfortunately, cannot be applied (or, extended) to more general dynamical systems (i.e., systems where $\underline{b}$ is time varying). For this reason we present an alternative approach to Friedland's derivation. We assume, at the onset, the existence of the decomposition for $\hat{\underline{x}}$ in Eq. (8); assume a specific structure for $\hat{\underline{b}}$; require that $\hat{\underline{x}}$ and $\hat{\underline{b}}$ be unbiased estimators of $\underline{x}$ and $\underline{b}$, respectively; and minimize the trace of the error covariance matrix of $\hat{\underline{x}}$ and $\hat{\underline{b}}$, with respect to a design matrix which appears in our assumed estimator of $\underline{b}$. We then demonstrate that the resulting $\hat{\underline{x}}$ and $\hat{\underline{b}}$ are indeed the minimum-variance estimators of $\underline{x}$

---

* $\underline{\xi} \perp \underline{\eta} \implies E\{\underline{\xi}\,\underline{\eta}'\} = 0$, always.

and $\underline{b}$; hence, our assumed decomposition of $\hat{\underline{x}}$ is a valid one.

Our derivation method, which is patterned after the Athans and Tse derivation of the Kalman filter (Ref. 13), is constructive in nature, which is important because it affords us the possibility of extending Friedland's results to more general classes of partitioned dynamical systems. Such extensions, which are beyond the scope of the present paper, will be described in a subsequent paper.

Some extensions of and work related to Friedland's decomposition can be found in Refs. 2, 3, 5, 6, 7, 8, 9, and 19. Agee and Turner (Ref. 5) augmented the bias states to the dynamical states and studied conditions under which $\hat{\underline{x}}$ can be decomposed as in Eq. (8). They showed that there are three basic requirements for a general linear filter to be decomposed as in Eq. (8): (1) V_x must satisfy a specific equation; (2) a gain condition of the form $K_{\hat{x}} = K_{\tilde{x}} + V_x K_{\hat{b}}$ must exist (the gains in this equation are related in an obvious way to estimates $\hat{\underline{x}}$, $\tilde{\underline{x}}$, and $\hat{\underline{b}}$); and (3) Eq. (8) must be true for all t (k) including t = 0 (k = 0). Lin and Sage (Ref. 2) formulate Friedland's problem in a different manner. In their approach, the bias states are thought of as unknown constant mean values of plant and measurement noise processes. The dynamical states are estimated as usual, whereas the unknown means are estimated using the maximum-likelihood formulation. Godbole (Ref. 3) has shown that Lin and Sage's results are identical to Friedland's. Bierman (Ref. 19) has obtained Friedland's decomposition using the square-root information filter, and, has extended Friedland's results to smoothing. Additionally, he gives matrix V_x the interesting and useful interpretation of a sensitivity matrix; i.e., $V_x = \text{grad} \, \hat{\underline{x}}$.
$$\underline{b}$$

In some applications, the bias states are used to model constant instrumentation error sources (Ref. 15), of which there can be a large number (e.g., 50). It can happen that not all of the error sources are significant, so that some of them should be deleted from the final filter;

or, that significant error sources may have been initially neglected, and should be included in the final filter. We are also interested in knowing whether or not it is possible to extend Friedland's decomposition to the situation where $\underline{b}$ is of variable dimension; but, in such a manner that all preceding calculations associated with obtaining $\hat{\underline{x}}(k)$ and $\tilde{\underline{x}}(k)$, for $\underline{b}$ of dimension r_1, do not have to be completely redone when $\underline{b}$ becomes of dimension r_2, where $r_2 > r_1$ (the case $r_2 < r_1$ is similar, but is not discussed in this paper). This extension is described in Section IV. Bierman (Ref. 19) has obtained similar results, using a square-root information filter approach.

Some important preliminary results are presented in Section II. A parallel development of both the discrete- and continuous-time decomposition, Eqs. (8) and (9), using our constructive approach, is given in Section III. The extension of Friedland's decomposition to the situation where $\underline{b}$ is of variable dimension is discussed in Section IV. Proofs for all theorems are given in appendices.

II. PRELIMINARY RESULTS

Recall that $\tilde{\underline{x}}$ in Eq. (8) is the bias-free estimate of $\underline{x}$, computed as if no biases were present. It is obtained by suppressing the bias states, $\underline{b}$, in $\mathcal{J}_0$ and applying the following Kalman filter to the resulting system:[*]

$$O(\tilde{\underline{x}}) = (A - \tilde{K}_x H)\,\tilde{\underline{x}} + \tilde{K}_x\,\underline{y}(\cdot\,|\,\mathcal{J}_0) \qquad (10)$$

For discrete-time systems, $\tilde{K}_x = \tilde{K}_x(k+1)$ and $H = H(k+1)$; matrix $\tilde{K}_x$ is the Kalman gain matrix for calculation of $\tilde{\underline{x}}$.

To facilitate our development of Eqs. (8) and (9), it is most convenient to introduce the following artificial system, $\mathcal{J}_2$:

---

[*] For discrete-time systems, change $(A - \tilde{K}_x H)$ in Eq. (10) to $(A - \tilde{K}_x HA)$.

-5-

$$\mathcal{A}_2 \begin{cases} O(\underline{x}_2) = A\,\underline{x}_2 + \underline{\xi} \quad ; \quad \underline{x}_2(0) & (11) \\ \underline{y}(\cdot \mid \mathcal{A}_2) = H\underline{x}_2 + \underline{\eta} & (12) \end{cases}$$

where $\underline{x}_2 \in R^n$ and all other quantities are as defined for $\mathcal{A}_0$. Only the actual measurement, $\underline{y}(\cdot \mid \mathcal{A}_0)$ exists; <u>artificial measurement</u> $\underline{y}(\cdot \mid \mathcal{A}_2)$ is non-existent, but, is useful for analysis purposes. Though we shall use $\underline{y}(\cdot \mid \mathcal{A}_2)$ in our analyses, it will not appear in any of our final results.

Let $\hat{\underline{x}}_2$ denote the estimate of artificial state vector $\underline{x}_2$ using the artificial measurement $\underline{y}(\cdot \mid \mathcal{A}_2)$; i.e.,[*]

$$O(\hat{\underline{x}}_2) = (A - \widetilde{K}_x H)\,\hat{\underline{x}}_2 + \widetilde{K}_x\,\underline{y}(\cdot \mid \mathcal{A}_2) \qquad (13)$$

Observe that the Kalman gain matrix for calculation of $\hat{\underline{x}}_2$ is exactly the same as for the calculation of $\hat{\underline{x}}$, since the well-known equations for the Kalman gain matrix (and associated Riccati equations) do not depend on the measurement process.

The following error processes and error covariance matrices will be used during our developments:

$$\underline{x}_e = \underline{x} - \hat{\underline{x}} \qquad (14a)$$

$$\underline{x}_{2e} = \underline{x}_2 - \hat{\underline{x}}_2 \qquad (14b)$$

$$\underline{b}_e = \underline{b} - \hat{\underline{b}} \qquad (14c)$$

$$\begin{pmatrix} P_x & P_{xb} \\ P_{bx} & P_b \end{pmatrix} = \begin{pmatrix} E\{\underline{x}_e\,\underline{x}_e'\} & E\{\underline{x}_e\,\underline{b}_e'\} \\ E\{\underline{b}_e\,\underline{x}_e'\} & E\{\underline{b}_e\,\underline{b}_e'\} \end{pmatrix} \qquad (15)$$

---

[*] For discrete-time systems, change $(A - \widetilde{K}_x H)$ in Eq. (13) to $(A - \widetilde{K}_x HA)$.

$$P_{x_2} = E\{\underline{x}_{2e}\,\underline{x}'_{2e}\}$$
(16a)

and

$$P_{bx_2} = E\{\underline{b}_e\,\underline{x}'_{2e}\}$$
(16b)

For discrete-time systems, all error covariances, as well as estimates are for filtered quantities (i.e., $\hat{\underline{x}} = \hat{\underline{x}}(k|k)$ and $P_x = P_x(k|k)$), unless otherwise noted.

The following theorems, which are proved in Appendix A, connect the states $\underline{x}$ and $\underline{x}_2$, and estimators $\tilde{\underline{x}}$ and $\hat{\underline{x}}_2$.

<u>Theorem 1.</u> If $\underline{x}(0) = \underline{x}_2(0)$ and Λ is $n \times r$, satisfying

$$O(\Lambda) = A\Lambda + B \ ; \qquad \Lambda(0) = 0 \ ,$$
(17)

then

$$\underline{x} = \underline{x}_2 + \Lambda \underline{b}$$
(18)

<u>Theorem 2.</u> If $\tilde{\underline{x}}(0) = \hat{\underline{x}}_2(0)$ and $\overline{\Lambda}$ is $n \times r$, satisfying (for continuous-time systems)

$$\dot{\overline{\Lambda}} = (A - \tilde{K}_x H)\,\overline{\Lambda} + \tilde{K}_x\,(H\Lambda + C)$$
(19a)

or (for discrete-time systems)

$$\overline{\Lambda}(k+1) = [I - \tilde{K}_x H]\,A\,\overline{\Lambda}(k) + \tilde{K}_x\,[H\Lambda(k+1) + C(k+1)]$$
(19b)

where $\overline{\Lambda}(0) = 0$, then

$$\tilde{\underline{x}} = \hat{\underline{x}}_2 + \overline{\Lambda}\,\underline{b}$$
(20)

Matrices Λ and $\overline{\Lambda}$ are connected, as we show in:

-7-

Theorem 3. Let $V_x \overset{\Delta}{=} \Lambda - \overline{\Lambda}$. Then V_x satisfies the following equation (for continuous-time systems)

$$\dot{V}_x = (A - \widetilde{K}_x H) V_x + (B - \widetilde{K}_x C) \tag{21a}$$

or (for discrete-time systems)

$$V_x(k+1) = [I - \widetilde{K}_x H][A V_x(k) + B] - \widetilde{K}_x C(k+1) \tag{21b}$$

where $V_x(0) = 0$.

Matrix V_x, in Theorem 3, is, as we show below, precisely the same matrix which appears in Eqs. (8) and (9).

III. DERIVATION OF FRIEDLAND'S DECOMPOSITION

A. Introduction

In this section, we <u>assume</u> that $\underline{x}$ and $\underline{b}$ can be estimated from the estimators

$$\hat{\underline{x}} = \widetilde{\underline{x}} + V_x \hat{\underline{b}} \tag{22}$$

and

$$O(\hat{\underline{b}}) = G_1 \hat{\underline{b}} + G_2 \underline{y}(\cdot \mid \mathscr{A}_0) + G_3 \widetilde{\underline{x}} \tag{23}$$

where $\widetilde{\underline{x}}$ is the bias-free estimate of $\underline{x}$, defined in Section II, and G_1, G_2, G_3 (which are matrices of appropriate dimensions) and V_x are determined such that: (1) $\hat{\underline{x}}$ and $\hat{\underline{b}}$ are unbiased estimates of $\underline{x}$ and $\underline{b}$, respectively, and (2) the trace of the error covariance matrix for $\hat{\underline{x}}$ and $\hat{\underline{b}}$ is minimized. We then show that the resulting estimators are <u>the</u> optimal estimators of $\underline{x}$ and $\underline{b}$.

B. Unbiasedness of the Estimators

The following two theorems demonstrate how G_1, G_3 and V_x must be chosen so that $\hat{\underline{x}}$ and $\hat{\underline{b}}$, given by Eqs. (22) and (23), respectively, are unbiased estimators of $\underline{x}$ and $\underline{b}$. Proofs of these theorems are given in Appendix B.

__Theorem 4.__ $\hat{\underline{b}}$ is an unbiased estimator of $\underline{b}$, iff

$$E\{\hat{\underline{b}}(0)\} = \underline{b} \tag{24}$$

and (for continuous-time systems)

$$G_1 = -G_2 (HV_x + C) \tag{25a}$$

$$G_3 = -G_2 H \tag{26a}$$

or (for discrete-time systems)

$$G_1(k+1) = I - G_2(k+1) S(k+1) \tag{25b}$$

$$G_3(k+1) = -G_2(k+1) HA \tag{26b}$$

where

$$S(k+1) = HAV_x + HB + C(k+1) \tag{27}$$

__Theorem 5.__ If $\hat{\underline{b}}$ is an unbiased estimator of $\underline{b}$ and V_x is as given in Eq. (21), then $\hat{\underline{x}}$ is an unbiased estimator of $\underline{x}$.

Theorems 4 and 5 show that if V_x is computed in a specific manner, and (for continuous-time systems)

$$\dot{\hat{\underline{b}}} = G_2 [\ y(t|\mathscr{A}_0) - (HV_x + C) \hat{\underline{b}} - H\tilde{\underline{x}}\ ] \tag{28a}$$

-9-

or (for discrete-time systems)

$$\hat{\underline{b}}(k+1) = [I - G_2(k+1) S(k+1)] \hat{\underline{b}}(k)$$

$$+ G_2(k+1) [\underline{y}(k|\mathcal{A}_0) - HA \tilde{\underline{x}}] , \qquad (28b)$$

then $\hat{\underline{x}}$ and $\hat{\underline{b}}$ are unbiased estimators of $\underline{x}$ and $\underline{b}$ in $\mathcal{A}_0$. Observe that V_x (which, as we pointed out in Section I, is a sensitivity matrix) maintains the unbiasedness of the estimator $\hat{\underline{x}}$, whereas G_1 and G_3, chosen as the preceding functions of G_2, maintain the unbiasedness of the estimator $\hat{\underline{b}}$. The only remaining matrix of design variables is G_2.

C. Optimal Gain Matrix, G_2

Matrix G_2 is chosen such that the trace of the error covariance matrix for the $\hat{\underline{x}}$ and $\hat{\underline{b}}$ estimators is minimized. That matrix, for $\mathcal{A}_0$, denoted $P(\cdot)$, has been defined in Eq. (15).

Theorem 6. If $P_{bx_2}(0) = 0$, then

$$P_{bx_2} = 0 \qquad \forall t \quad \text{or} \quad \forall k \qquad (29)$$

$$P_x = P_{x_2} + V_x P_b V'_x \qquad (30)$$

and*

$$P_{xb} = V_x P_b , \qquad (31)$$

so that

---

*Agee and Turner (Ref. 5) show that their gain condition holds if Eq. (31) is satisfied. A gain condition, comparable to theirs, does not fall out of our derivation.

-10-

$$P(\cdot) = \left(\begin{array}{c|c} P_{x_2} & 0 \\ \hline 0 & 0 \end{array} \right) + \left(\frac{V_x}{I} \right) P_b \left(\frac{V_x}{I} \right)' \qquad (32)$$

The proof of this theorem, as well as the other theorems in this paragraph, are given in Appendix C. Matrices V_x and $^* P_{x_2}$ are not functions of G_2; hence, <u>minimizing tr $P(\cdot)$ with respect to G_2 is equivalent to minimizing tr $V_x P_b V_x'$, and subsequently tr P_b with respect to G_2.</u>

<u>Theorem 7.</u> tr P_b is minimum with respect to G_2 iff, (for continuous-time systems)

$$G_2^* = P_b (H V_x + C)' R^{-1} \qquad (33a)$$

or (for discrete-time systems)

$$G_2^*(k+1) = P_b S'(k+1) \left[S(k+1) P_b S'(k+1) + H P_{x_2}(k+1|k) H' + R(k+1) \right]^{-1} \quad (33b)$$

where

$$P_{x_2}(k+1|k) = A P_{x_2}(k|k) A' + Q . \qquad (34)$$

Additionally, the optimal error covariance matrix for P_b can be computed from (for continuous-time systems)

$$\dot{P}_b = -P_b (H V_x + C)' R^{-1} (H V_x + C) P_b \qquad (35a)$$

or (for discrete-time systems)

---

*Matrix P_{x_2}, defined in Eq. (16a), is associated with $\hat{\underline{x}}_2$ for $\mathscr{A}_2$.

$$P_b(k+1) = [I - G_2^*(k+1) S(k+1)] P_b \qquad (35b)$$

D. Summary

It is helpful, at this point, to summarize our results. We let $G_2^* = K_b$. For <u>continuous-time systems</u> we have shown that

$$\dot{\hat{\underline{b}}} = -K_b (HV_x + C) \hat{\underline{b}} + K_b \tilde{\underline{y}}(t | \mathscr{A}_0) \qquad (36)$$

where

$$\tilde{\underline{y}}(t | \mathscr{A}_0) = \underline{y}(t | \mathscr{A}_0) - H\tilde{\underline{x}} \qquad (37)$$

$$K_b(t) = P_b (HV_x + C)' R^{-1} \qquad (38)$$

$$\dot{P}_b = -P_b(HV_x + C)' R^{-1} (HV_x + C) P_b \; ; \quad P_b(0) \qquad (39)$$

and $V_x(t)$ is computed via Eq. (21a). For <u>discrete-time systems</u> we have shown that

$$\hat{\underline{b}}(k+1) = [I - K_b(k+1) S(k+1)]\hat{\underline{b}}(k) + K_b(k+1) \tilde{\underline{y}}(k+1 | \mathscr{A}_0) \qquad (40)$$

where

$$\tilde{\underline{y}}(k+1 | \mathscr{A}_0) = \underline{y}(k+1 | \mathscr{A}_0) - HA\tilde{\underline{x}} \qquad (41)$$

$$K_b(k+1) = P_b S'(k+1)[S(k+1) P_b S'(k+1) + H P_{x_2}(k+1|k) H' + R(k+1)]^{-1} \qquad (42)$$

$$P_b(k+1) = [I - K_b(k+1) S(k+1)] P_b(k) \; ; \quad P_b(0) \qquad (43)$$

-12-

and $S(k+1)$ is computed via Eq. (27), which uses $V_x(k)$ from Eq. (21b). In both cases, $\hat{\underline{x}}$ is computed via Eq. (22).

Observe that the bias-free estimator, $\tilde{\underline{x}}$, couples into the calculation of $\hat{\underline{b}}$, through calculations of $\tilde{\underline{y}}(\cdot \mid \mathcal{J}_0)$ and V_x; but, $\hat{\underline{b}}$ in no way affects $\tilde{\underline{x}}$.

E. <u>Estimators $\hat{\underline{x}}$ and $\hat{\underline{b}}$ are the Minimum-Variance Estimators</u>

Kalman (Refs. 10 and 11) has shown that a necessary and sufficient condition for an estimator ($\hat{\underline{b}}$ or $\hat{\underline{x}}$) to be the optimal (i.e., minimum-variance) estimator is, that (for continuous-time systems)

$$ E\left\{ \begin{pmatrix} \underline{x}_e \\ \underline{b}_e \end{pmatrix} \underline{y}'(\tau \mid \mathcal{J}_0) \right\} = 0 \ , \qquad 0 \le \tau \le t \qquad (44a) $$

or (for discrete-time systems)

$$ E\left\{ \begin{pmatrix} \underline{x}_e \\ \underline{b}_e \end{pmatrix} \underline{y}'(i \mid \mathcal{J}_0) \right\} = 0 \ , \qquad i=0,1,2,\ldots,k \qquad (44b) $$

In Appendix D, we use these criteria to prove:

<u>Theorem 8.</u> Our optimal linear estimators of $\hat{\underline{x}}$ and $\hat{\underline{b}}$, given by Eqs. (22) and (for continuous-time systems) (36) - (39), or (for discrete-time systems) (40) - (43), are <u>the</u> optimal estimators of $\underline{x}$ and $\underline{b}$ in $\mathcal{J}_0$.

With this result we have completed our constructive derivation of Friedland's decomposition.

IV. MULTISTAGE ESTIMATION OF BIASES

We now direct our attention at the problem of adding (or deleting) bias states. We wish to do this in such a manner that previously computed quantities can be used to obtain the new estimates of $\underline{x}$ and $\underline{b}$ in $\mathscr{A}_0$. Let $\underline{b} = \text{col}\,(\underline{b}_1, \underline{b}_2)$, where $\underline{b}_1 \in R^{r_1}$ and $\underline{b}_2 \in R^{r_2}$. Suppose that, initially, $\underline{b} = \underline{b}_1$; then $\mathscr{A}_0$ reduces to $\mathscr{A}_1$, where*

$$\mathscr{A}_1 \begin{cases} \underline{x}_1(k+1) = A\,\underline{x}_1(k) + B_1\,\underline{b}_1 + \underline{\xi} \quad ; \quad \underline{x}_1(0) & (45) \\[2ex] \underline{b}_1(k+1) = \underline{b}_1(k) = \underline{b}_1 & (46) \\[2ex] \underline{y}(k|\mathscr{A}_1) = H\,\underline{x}_1(k) + C_1\,\underline{b}_1 + \underline{\eta} & (47) \end{cases}$$

All quantities in $\mathscr{A}_1$ are analogous to those in $\mathscr{A}_0$. When we believe that $\underline{b} = \underline{b}_1$, then $\underline{y}(k|\mathscr{A}_1)$ is the actual measurement.

Let us assume that $\hat{\underline{x}}_1$ has been computed for $\mathscr{A}_1$, as

$$\hat{\underline{x}}_1 = \tilde{\underline{x}}_1 + V_{x_1}\,\hat{\underline{b}}_1^{\,*} \,, \tag{48}$$

where $\tilde{\underline{x}}_1$ is the bias-free estimate of $\underline{x}_1$, and $\hat{\underline{b}}_1^{\,*}$ and V_{x_1} are computed from our results in Section III, with the appropriate substitutions, that $B = B_1$ and $C = C_1$. Now, let us add additional bias states, $\underline{b}_2$, so that our system is $\mathscr{A}_0$, with $\underline{b} = \text{col}\,(\underline{b}_1, \underline{b}_2)$, $B = (B_1 \mid B_2)$, and $C = (C_1 \mid C_2)$. We wish to estimate $\underline{b}$ in $\mathscr{A}_0$ using the already computed $\hat{\underline{b}}_1^{\,*}$, and, we wish to compute $\hat{\underline{x}}$ by means of Eq. (22) using $\hat{\underline{b}}_1^{\,*}$ and the already computed $\tilde{\underline{x}}_1$.

---

*This part of our paper, which deals with recursive bias estimation, is presented only for discrete-time systems, since it seems to be most practical for such systems.

The key to solving this problem is to recognize that $\overset{\bullet}{\underline{b}}$, given by Eqs. (40) - (43), can be associated with the following <u>auxiliary parameter estimation problem</u> (Ref. 16): Obtain minimum-variance estimates of the constant parameter vector, $\underline{b}$, from the (auxiliary) measurement equation

$$\widetilde{\underline{r}}(k+1) = S(k+1)\,\underline{b} + \overline{\underline{v}}(k+1) \tag{49}$$

where $\overline{\underline{v}}(k+1)$ is, by definition, a zero-mean white noise sequence, and

$$E\{\,\overline{\underline{v}}(k+1)\,\overline{\underline{v}}'(k+1)\} \overset{\Delta}{=} R(k+1) + H(k+1)\,P_{x_2}(k+1|k)\,H'(k+1) \tag{50}$$

Actually, $\widetilde{\underline{r}}(k+1) = \widetilde{\underline{y}}(k+1|\mathscr{A}_0)$, and, we see from Eq. (41) that $\widetilde{\underline{r}}(k+1)$ can be assembled from the actual measurement $\underline{y}(k+1|\mathscr{A}_0)$ and the bias-free estimate of $\underline{x}$, $\widetilde{\underline{x}}$; i.e.,

$$\widetilde{\underline{r}}(k+1) = \underline{y}(k+1|\mathscr{A}_0) - HA\widetilde{\underline{x}} \tag{51}$$

Numerous authors (Refs. 4, 8, and 14, for example) have developed algorithms for estimation of a constant parameter vector which are recursive in the dimension of that vector, as well as sequential in time. Mendel (Theorem 3, Ref. 4), for example, shows that

$$\hat{\underline{b}}_1(k+1) = \hat{\underline{b}}_1^*(k+1) - G(k+1)\,\hat{\underline{b}}_2(k+1) \tag{52a}$$

and

$$\overset{\bullet}{\underline{b}}_2(k+1) = C(k+1)\big[\,\underline{u}_2(k+1) - G'_{12}(k+1)\,\hat{\underline{b}}_1^*(k+1)\big] \tag{52b}$$

where[*] $G(k+1)$, $C(k+1)$, $\underline{u}_2(k+1)$, and $G_{12}(k+1)$ are computed from other

_____

[*] Our statement of Eqs. (52a) and (52b) retains Mendel's notation; matrix $C(k+1)$ in (52b) is <u>not</u> the same as our matrix C.

-15-

equations given by Mendel. These equations are not needed for our purposes; hence, they are not given.

We conclude, therefore, that it is indeed possible to estimate $\underline{b}$ in $\mathcal{A}_0$ using the already computed $\hat{\underline{b}}_1^*$. This is accomplished by applying a parameter estimation algorithm, which is recursive in the dimension of $\underline{b}$ and sequential in time, to estimation of $\underline{b}$ for the auxiliary measurement $\widetilde{\underline{r}}(k+1)$.

The following theorem demonstrates how to compute $\hat{\underline{x}}(k)$ from previously computed $\widetilde{\underline{x}}_1(k)$ and $\hat{\underline{b}}_1^*(k)$, and, from $\hat{\underline{b}}_2(k)$. These results are similar to those in Bierman (Ref. 19).

<u>Theorem 9.</u> Given $\widetilde{\underline{x}}_1(k)$ and $\hat{\underline{b}}_1^*(k)$, for $\mathcal{A}_1$. State estimate $\hat{\underline{x}}(k)$, for $\mathcal{A}_0$, can be computed from:

$$\hat{\underline{x}}(k) = \widetilde{\underline{x}}(k) + V_{x_1}(k)\,\hat{\underline{b}}_1^*(k) + \left[V_{x_2}(k) - V_{x_1}(k)G(k) \right]\hat{\underline{b}}_2(k) \tag{53}$$

where, for $i = 1, 2$,

$$V_{x_i}(k+1) = \left[I - \widetilde{K}_x(k+1)\,H(k+1) \right]\left[A\,V_{x_i}(k) + B_i \right]$$

$$- \widetilde{K}_x(k+1)\,C_i(k+1) \quad ; \quad V_{x_i}(0) = 0 , \tag{54}$$

and, $\hat{\underline{b}}_2(k)$ and $G(k)$ are computed from an algorithm which is recursive in the dimension of $\underline{b}$ and is sequential in time, such as the one in Mendel (Theorem 3, Ref. 4).

The proof of this theorem is given in Appendix E. It is important to understand that $\underline{\widetilde{x}}(k) = \widetilde{\underline{x}}_1(k)$; i.e., adding (or deleting) biases does not affect the "bias-free estimates." We see, from Eq. (53), that $\hat{\underline{x}}(k)$ does not have to be completely recomputed when biases are added. In fact, another way to write Eq. (53) is, as

-16-

$$\hat{\underline{x}}(k) = \hat{\underline{x}}_1(k) + \left[V_{x_2}(k) - V_{x_1}(k) G(k) \right] \hat{\underline{b}}_2(k) \tag{55}$$

which makes this even more evident.

V. CONCLUSIONS

We have rederived Friedland's decomposition, Eq. (8), by means of a constructive approach, which lends itself to studying extensions of his results to more complex systems, where biases are themselves modeled as first-order Markov processes (Ref. 17, for example). These extensions have been made and will be reported on in a future publication. Our constructive approach gives some insight and physical meaning into certain quantities, such as V_x, which appeared in Friedland's solution. The key to our constructive approach is the artificial system, $\mathcal{S}_2$; for, through that system, we are able to make connections between $\underline{x}$ and $\underline{x}_2$ as well as $\tilde{\underline{x}}$ and $\hat{\underline{x}}_2$.

As a consequence of our approach, we showed that $\hat{\underline{b}}$ can be given the interpretation of a minimum-variance parameter estimator which is associated with an auxiliary measurement. With this interpretation, we can use well-known algorithms for estimating the bias which are recursive in the dimension of the bias as well as sequential in time. Hence, we have shown that it is possible to add (or delete) bias states in such a manner that previously computed quantities can be used to obtain new estimates of $\underline{x}$ and $\underline{b}$.

-17-

APPENDIX A.   PROOFS OF PRELIMINARY RESULTS

The proofs presented in this and subsequent appendices are for the continuous-time case. It is felt that these proofs give the essence of the constructive method in a clearer fashion than those of the discrete-time case. Many of the discrete-time proofs are inductive in nature.

1. Proof of Theorem 1

From Eqs. (11) and (17), it follows that

$$\frac{d}{dt}(\underline{x}_2 + \Lambda \underline{b}) = \dot{\underline{x}}_2 + \dot{\Lambda} \underline{b} = A(\underline{x}_2 + \Lambda \underline{b}) + B \underline{b} + \underline{\xi} \; ; \qquad (A-1)$$

but, $\underline{x}_2(0) + \Lambda(0)\underline{b} = \underline{x}_2(0) = \underline{x}(0)$; hence, $\underline{x}_2 + \Lambda \underline{b}$ satisfies the same differential equation (d.e.) as $\underline{x}$. By uniqueness, therefore, $\underline{x}_2 + \Lambda \underline{b} = \underline{x}$, which is Eq. (18).

2. Proof of Theorem 2

From Eqs. (13) and (19a), it follows that

$$\frac{d}{dt}(\hat{\underline{x}}_2 + \overline{\Lambda}\underline{b}) = \dot{\hat{\underline{x}}}_2 + \dot{\overline{\Lambda}}\underline{b}$$

$$= (A - \widetilde{K}_x H)(\hat{\underline{x}}_2 + \overline{\Lambda}\underline{b}) + \widetilde{K}_x [\underline{y}(t|\mathscr{A}_2) + H\Lambda\underline{b} + C\underline{b}] \quad (A-2)$$

Additionally, from Eqs. (3), (12), and (18), we see that

$$\underline{y}(t|\mathscr{A}_0) = \underline{y}(t|\mathscr{A}_2) + H\Lambda \underline{b} + C \underline{b} \qquad (A-3)$$

so that Eq. (A-2) can also be written, in terms of $\underline{y}(t|\mathscr{A}_0)$, as

-18-

$$\frac{d}{dt} (\hat{\underline{x}}_2 + \overline{\Lambda} \, \underline{b}) = (A - \widetilde{K}_x H) (\hat{\underline{x}}_2 + \overline{\Lambda} \, \underline{b}) + \widetilde{K}_x \, \underline{y}(t \,|\, \mathscr{I}_0) ; \qquad (A-4)$$

but, $\hat{\underline{x}}_2(0) + \overline{\Lambda}(0) \, \underline{b} = \hat{\underline{x}}_2(0) = \widetilde{\underline{x}}(0)$; hence, $\hat{\underline{x}}_2 + \overline{\Lambda} \, \underline{b}$ satisfies the same d. e. as $\widetilde{\underline{x}}$. By uniqueness, therefore, $\hat{\underline{x}}_2 + \overline{\Lambda} \, \underline{b} = \widetilde{\underline{x}}$, which is Eq. (20).

3. Proof of Theorem 3

To begin, observe from Theorems 1 and 2 that $V_x(0) = \Lambda(0) - \overline{\Lambda}(0) = 0$. Then, using Eqs. (17) and (19a), we see that

$$\dot{V}_x = \dot{\Lambda} - \dot{\overline{\Lambda}} = (A - \widetilde{K}_x H) V_x + (B - \widetilde{K}_x C) , \qquad (A-5)$$

which is Eq. (21a).

APPENDIX B. UNBIASEDNESS PROOFS

1. Proof of Theorem 4

From Eqs. (2a) and (23), we see that

$$\frac{d}{dt}(\underline{b} - \hat{\underline{b}}) = -G_1 \hat{\underline{b}} - G_2 \underline{y}(t|\mathcal{A}_0) - G_3 \tilde{\underline{x}} \tag{B-1}$$

Substitute Eqs. (3), (18) and (20) into Eq. (B-1), take the expected value of the resulting equation, and use the fact that $\hat{\underline{x}}_2$ is an unbiased estimator of $\underline{x}_2$, to show that

$$\frac{d}{dt} E\{\underline{b} - \hat{\underline{b}}\} = -G_1 E\{\hat{\underline{b}}\} - (G_2 H\Lambda + G_2 C + G_3 \overline{\Lambda}) \underline{b}$$

$$-(G_2 H + G_3) E\{\underline{x}_2\} \tag{B-2}$$

To prove __sufficiency__, assume $E\{\hat{\underline{b}}\} = E\{\underline{b}\}$; then, from Eq. (B-2),

$$G_1 + G_2 H\Lambda + G_2 C + G_3 \overline{\Lambda} = 0 \tag{B-3}$$

and

$$G_2 H + G_3 = 0 \tag{B-4}$$

Solve Eq. (B-4) for G_3 to obtain Eq. (26a); substitute that expression into Eq. (B-3); and use the fact, from Theorem 3, that $V_x = \Lambda - \overline{\Lambda}$, to show that

$$G_1 = -G_2 (H V_x + C) , \tag{B-5}$$

which is Eq. (25a).

To prove __necessity__, substitute Eqs. (25a) and (26a) into Eq. (B-2), to see that

-20-

$$\frac{d}{dt} E\{ \underline{b} - \hat{\underline{b}} \} = G_1 E\{ \underline{b} - \hat{\underline{b}} \} \qquad \text{(B-6)}$$

Because of Eq. (24), the unique solution to this equation is $E\{ \underline{b} - \hat{\underline{b}} \} = \underline{0}$; hence, $\hat{\underline{b}}$ is an unbiased estimate of $\underline{b}$.

2. Proof of Theorem 5

From Eqs. (18), (8), and (20), and Theorem 3, we see that

$$\underline{x} - \hat{\underline{x}} = (\underline{x}_2 - \hat{\underline{x}}_2) + V_x (\underline{b} - \hat{\underline{b}}) \qquad \text{(B-7)}$$

Now, $\hat{\underline{x}}_2$ is the optimal (unbiased) estimate of $\underline{x}_2$, so that $E\{ \underline{x}_2 - \hat{\underline{x}}_2 \} = \underline{0}$; hence, $E\{ \underline{x} - \hat{\underline{x}} \} = \underline{0}$, which proves that $\hat{\underline{x}}$ is an unbiased estimate of $\underline{x}$.

APPENDIX C. OPTIMIZATION PROOFS

1. Proof of Theorem 6

To begin, we show that if $P_{bx_2}(0) = 0$, then $P_{bx_2} = 0$ for all t. We obtain a differential equation for P_{bx_2} by differentiating Eq. (16b),

$$\dot{P}_{bx_2} = E\{\underline{\dot{b}}_e \underline{x}'_{2e}\} + E\{\underline{b}_e \underline{\dot{x}}'_{2e}\} \tag{C-1}$$

From well-known filter results (Ref. 20, for example), we know that

$$\underline{\dot{x}}_{2e} = (A - \widetilde{K}_x H)\underline{x}_{2e} + \underline{\xi} - \widetilde{K}_x \underline{\eta} \tag{C-2}$$

Additionally, from Eqs. (14c), (2a), and (28a), we find that

$$\underline{\dot{b}}_e = -G_2 [\, (HV_x + C)\underline{b}_e + H\underline{x}_{2e} + \underline{\eta}\,] \tag{C-3}$$

It is then relatively straightforward to show that

$$E\{\underline{\dot{b}}_e \underline{x}'_{2e}\} = -G_2(HV_x + C) P_{bx_2} - G_2 H P_{x_2} + \frac{1}{2} G_2 R \widetilde{K}'_x \tag{C-4}$$

and

$$E\{\underline{b}_e \underline{\dot{x}}'_{2e}\} = P_{bx_2}(A - \widetilde{K}_x H)' + \frac{1}{2} G_2 R \widetilde{K}'_x \tag{C-5}$$

Substituting Eqs. (C-4) and (C-5) into Eq. (C-1), making use of the fact that $H P_{x_2} = R\widetilde{K}'_x$, we see that P_{bx_2} satisfies the following linear homogeneous d. e.:

$$\dot{P}_{bx_2} = -G_2(HV_x + C) P_{bx_2} + P_{bx_2}(A - \tilde{K}_x H) \tag{C-6}$$

Since $P_{bx_2}(0) = 0$, it follows that $P_{bx_2} = 0$, all t.

Next, we develop the expression for P_x in Eq. (30). From Eq. (B-7), we see that

$$P_x = P_{x_2} + V_x P_{bx_2} + P_{x_2b} V'_x + V_x P_b V'_x \ , \tag{C-7}$$

and this expression reduces to Eq. (30), since $P_{bx_2} = 0$.

In a similar manner, it follows from Eqs. (B-7) and (29), that

$$P_{xb} = P_{bx_2} + V_x P_b = V_x P_b \tag{C-8}$$

which is Eq. (31).

Finally, to obtain Eq. (32) for P(t), substitute Eqs. (30) and (31) into the right-hand side of Eq. (15) and collect the terms as shown.

2. Proof of Theorem 7

We seek to minimize tr P_b with respect to G_2, where

$$\dot{P}_b = -G_2(HV_x + C) P_b - P_b(HV_x + C)' G'_2 + G_2 R G'_2 \tag{C-9}$$

Equation (C-9) follows, in the usual manner, from Eqs. (C-3) and (29). Our approach is the matrix minimum principle, in the manner of Ref. 12, applied to the following Hamiltonian, H:

$$H = \text{tr} \left[\dot{P}_b(t)\, T(t) \right] \qquad\qquad (C\text{-}10)$$

From the condition that $\partial H / \partial G_2 = 0$, and gradient matrix operations, we see that $\text{tr}\, P_b$ is minimum with respect to G_2, if

$$G_2 = P_b\,(HV_x + C)'\, R^{-1} \;, \qquad\qquad (C\text{-}11)$$

which is Eq. (33a).

Equation (35a) follows upon substitution of Eq. (C-11) into Eq. (C-9).

APPENDIX D. OPTIMALITY PROOF

To prove Theorem 8, we use the orthogonality principle (Refs. 10 and 11) stated in Eq. (44a). Because the details of the proof are somewhat intricate, and do not contribute to further understanding of this paper's problems, we give a brief sketch of the proof.   Details can be found in the second author's Ph. D. dissertation (Ref. 18).

From Eq. (B-7), we see that

$$E\left\{\begin{pmatrix}\underline{x}_e \\ \underline{b}_e\end{pmatrix} \underline{y}'(\tau\,|\,\mathcal{A}_0)\right\} = \begin{pmatrix}E\{\underline{x}_{2e}\,\underline{y}'(\tau\,|\,\mathcal{A}_0)\} \\ 0\end{pmatrix} + \begin{pmatrix}V_x \\ I\end{pmatrix} E\{\underline{b}_e\,\underline{y}'(\tau\,|\,\mathcal{A}_0)\}$$

$$(D-1)$$

Using Eq. (A-3) and the fact that $\hat{\underline{x}}_2$ is the optimal estimate for $\underline{y}(t\,|\,\mathcal{A}_2)$, it follows that $E\{\underline{x}_{2e}\,\underline{y}'(\tau\,|\,\mathcal{A}_0)\} = 0$, all $0 \leq \tau \leq t$. To show that $E\{\underline{b}_e\,\underline{y}'(\tau\,|\,\mathcal{A}_0)\} = 0$, all $0 \leq \tau \leq t$, consider that expectation first for $0 \leq \tau < t$ and then for $\tau = t$. In the former situation, show from Eq. (C-3) that $E\{\dot{\underline{b}}_e\,\underline{y}'(\tau\,|\,\mathcal{A}_0)\}$ satisfies a linear homogeneous d. e. with zero initial conditions; hence, $E\{\underline{b}_e\,\underline{y}'(\tau\,|\,\mathcal{A}_0)\} = \underline{0}$, for $0 \leq \tau < t$. For the latter situation, use the fact that $\hat{\underline{b}}$ is the optimal estimate for observation equation (37); hence, $E\{\underline{b}_e(t)\,\tilde{\underline{y}}'(t\,|\,\mathcal{A}_0)\} = 0$, so that

$$E\{\underline{b}_e(t)\,\tilde{\underline{y}}'(t\,|\,\mathcal{A}_0)\} = E\{\underline{b}_e\,\tilde{\underline{x}}'(t)\}\,H'(t) \qquad (D-2)$$

Finally, express $\tilde{\underline{x}}(t)$ as a linear transformation of the measurements over the interval $[0, t]$ and show that the right-hand side of Eq. (D-2) is zero. By these arguments, we have shown that our results give the optimal linear estimators of $\underline{x}$ and $\underline{b}$. Because all noise processes are gaussian, it then follows that our results give the optimal estimators of $\underline{x}$ and $\underline{b}$ (Ref. 20).

-25-

APPENDIX E.   PROOF OF MULTISTAGE RESULTS

It is straightforward to show, from Eq. (54), that

$$V_x = \left(V_{x_1} \mid V_{x_2} \right),$$ (E-1)

where V_x is defined in Eq. (21b). From Eqs. (8) and (E-1), it follows that

$$\hat{\underline{x}} = \tilde{\underline{x}} + V_{x_1} \hat{\underline{b}}_1 + V_{x_2} \hat{\underline{b}}_2 ,$$ (E-2)

and, substituting Eq. (52a) into Eq. (E-2), we obtain Eq. (55).

REFERENCES

1. B. Friedland, "Treatment of Bias in Recursive Filtering," _IEEE Trans. on Automatic Control_, Vol. AC-14, August 1969, pp. 359-367.

2. J. L. Lin and A. P. Sage, "Algorithms for Discrete Sequential Maximum Likelihood Bias Estimation and Associated Error Analysis," _IEEE Trans. on Systems, Man, and Cybernetics_, Vol. SMC-1, October 1974, pp. 314-324.

3. S. S. Godbole, "Comparison of Friedland's and Lin-Sage's Bias Estimation Algorithms," _IEEE Trans. on Automatic Control_, Vol. AC-19, April 1974, pg. 143.

4. J. M. Mendel, "Multistage Least-Squares Estimators," _IEEE Trans. on Automatic Control_, Vol. AC-20, December 1975, pp. 775-782.

5. W. S. Agee and R. H. Turner, "Optimal Estimation of Measurement Bias," Report AD-753961, White Sands Missile Range, New Mexico, December 1972.

6. J. M. Mendel, "Extension of Friedland's Bias Filtering Techniques to a Class of Nonlinear Systems," _IEEE Trans. on Automatic Control_, Vol. AC-21, April 1976, pp. 296-298.

7. E. C. Tacker and C. C. Lee, "Linear Filtering in the Presence of Time-Varying Bias," _IEEE Trans. on Automatic Control_, Vol. AC-17, December 1972, pp. 828-829.

8. G. J. Bierman, "Measurement Updating Using the U-D Factorization," Proceedings of the 1975 IEEE Conf. on Decision and Control, Houston, Texas, December 1975, pp. 337-346.

9. E. L. Shreve and W. R. Hedrick, "Separating Bias and State Estimates in a Recursive Second-Order Filter," _IEEE Trans. on Automatic Control_, Vol. AC-19, October 1974, pp. 585-586.

10. R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," _Trans. ASME, Series D: J. of Basic Engineering_, Vol. 82, pp. 35-45, 1960.

11. R. E. Kalman and R. S. Bucy, "New Results in Linear Filtering and Prediction Theory," _Trans. ASME, Series D: J. of Basic Engineering_, Vol. 83, pp. 95-108, 1961.

12. M. Athans, "The Matrix Minimum Principle," _Information and Control_, Vol. 11, 1968, pp. 592-606.

13. M. Athans and E. Tse, "A Direct Derivation of the Optimal Linear Filter Using the Maximum Principle," *IEEE Trans. on Automatic Control*, Vol. AC-12, No. 6, December 1967, pp. 690-698.

14. C. L. Lawson and R. J. Hanson, "Solving Least Squares Problems," Prentice-Hall, New Jersey, 1974.

15. R. Barton and A. Ljungwe, "Trajectory Reconstruction Utilizing Kalman Filter Estimation Theory," AIAA Paper 71-933, presented at 1971 AIAA Guidance, Control and Flight Mechanic Conf., Hempstead, New York, August 1971.

16. J. M. Mendel, "Discrete Techniques of Parameter Estimation," Marcel Dekker, New York, 1973.

17. A. Tanaka, "Parallel Computation of Linear Discrete Filtering," *IEEE Trans. on Automatic Control*, Vol. AC-20, pp. 573-575, August 1975.

18. H. D. Washburn, Ph.D. Dissertation, to appear in 1977.

19. G. J. Bierman, "The Treatment of Bias in the Square-Root Information Filter/Smoother," *Journal of Optimization Theory and Applications*, Vol. 16, July 1975, pp. 165-178.

20. J. S. Meditch, "Stochastic Optimal Linear Estimation and Control," McGraw-Hill, New York, 1969.